Tohzto 2008760

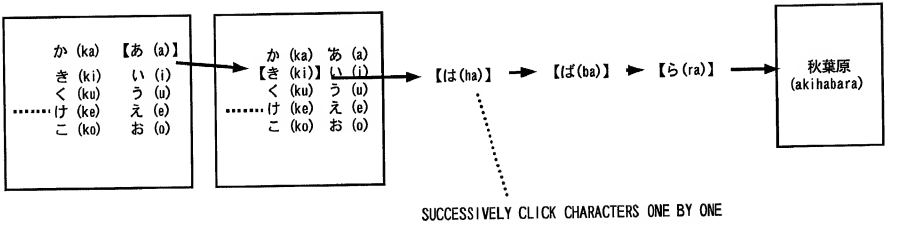


FIG. 1

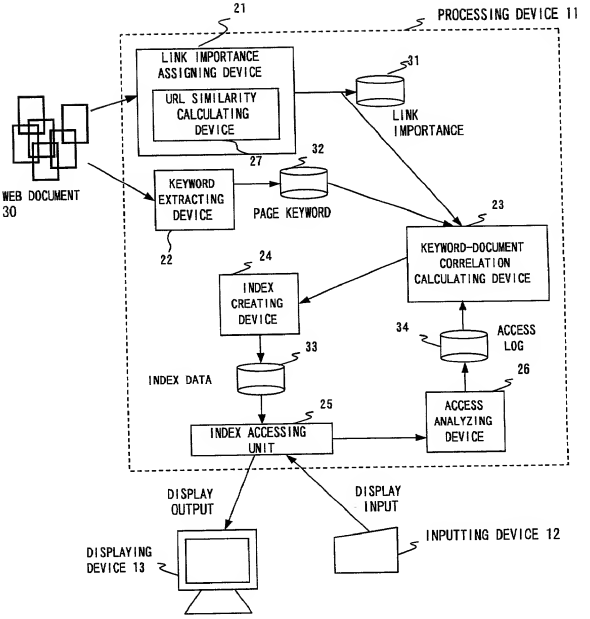

21
PROCESSING DEVICE 11
LINK IMPORTANCE ASSIGNING DEVICE
URL SIMILARITY CALCULATING DEVICE
27
31
LINK IMPORTANCE
32
KEYWORD EXTRACTING DEVICE
PAGE KEYWORD
WEB DOCUMENT 30
22
23
KEYWORD-DOCUMENT CORRELATION CALCULATING DEVICE
24
INDEX CREATING DEVICE
34
ACCESS LOG
33
INDEX DATA
26
ACCESS ANALYZING DEVICE
25
INDEX ACCESSING UNIT
DISPLAY OUTPUT
DISPLAY INPUT
INPUTTING DEVICE 12
DISPLAYING DEVICE 13

FIG. 2

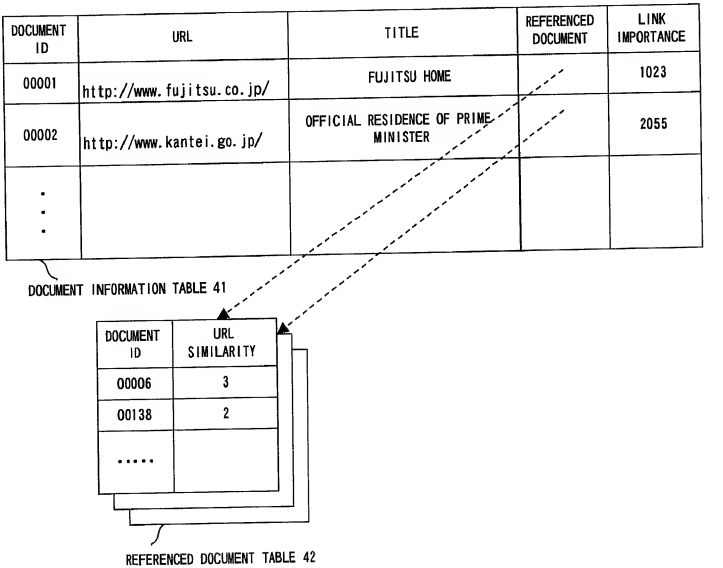

DOCUMENT ID
URL
TITLE
REFERENCED DOCUMENT
LINK IMPORTANCE
00001
http://www.fujitsu.co.jp/
FUJITSU HOME
1023
00002
http://www.kantei.go.jp/
OFFICIAL RESIDENCE OF PRIME MINISTER
2055
DOCUMENT INFORMATION TABLE 41
DOCUMENT ID
URL SIMILARITY
00006
3
00138
2
.....
REFERENCED DOCUMENT TABLE 42

FIG. 3

JP2002259560

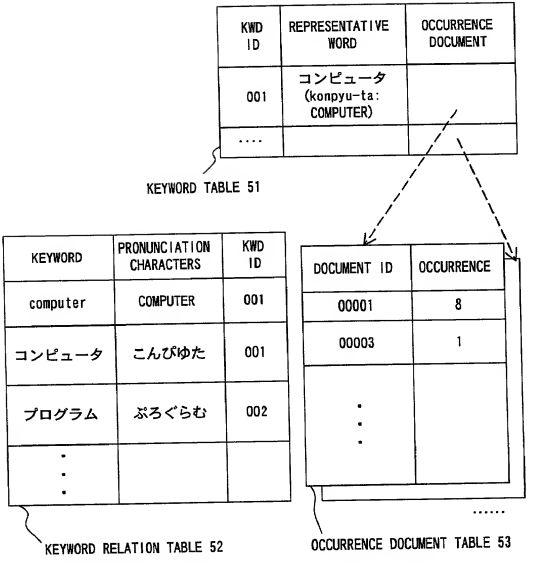


FIG. 4

CHARACTER STRING	FOLLOWING CHARACTER STRINGS	KEYWORD STRINGS
TOP	あ (a), い (i), ..	
あ (a)	あいぼ(aibo), あお(ao), ...	
あいぼ (aibo)		相棒 (aibou:MATE), アイボリー(aibori : IVORY).
あお (ao)	あおぞ(aozora)	青 (ao: BLUE), 蒼 (ao: DARK BLUE), ...

INDEX INFORMATION TABLE 61

KEYWORD ID	CORRELATED DOCUMENT ID STRINGS
093 321	0005, 0008, .. 0004, 0008, ...

CORRELATED DOCUMENT TABLE 62

DOCUMENT ID	CORRELATED KEYWORD ID STRINGS
0005 0008 . . .	093, 099, 122, ... 093, 156, 321, ...

CORRELATED KEYWORD TABLE 63

FIG. 5

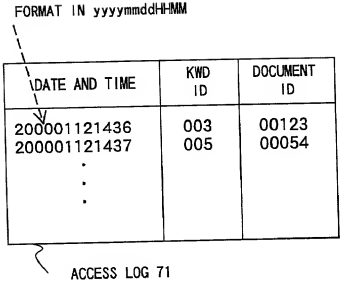

FORMAT IN yyyymmddHHMM
DATE AND TIME
KWD ID
DOCUMENT ID
200001121436
200001121437
003
005
00123
00054
ACCESS LOG 71

FIG. 6

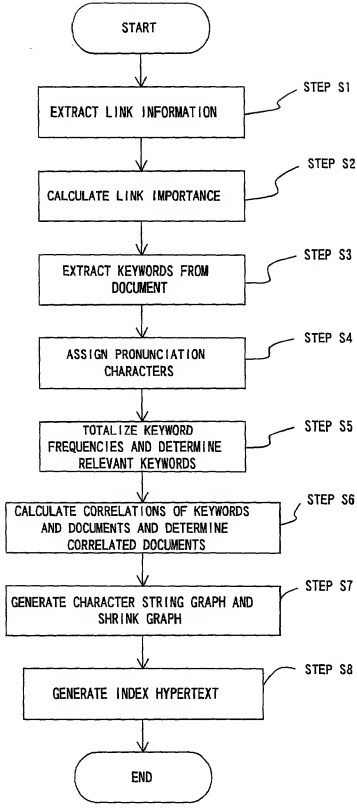
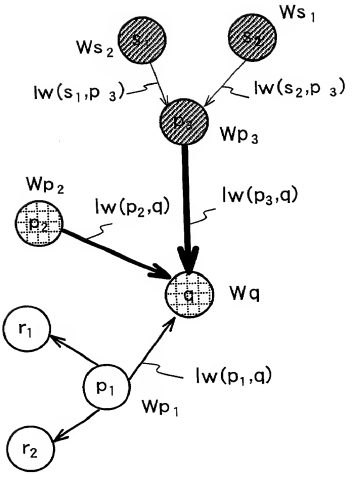

START
EXTRACT LINK INFORMATION
STEP S1
CALCULATE LINK IMPORTANCE
STEP S2
EXTRACT KEYWORDS FROM DOCUMENT
STEP S3
ASSIGN PRONUNCIATION CHARACTERS
STEP S4
TOTALIZE KEYWORD FREQUENCIES AND DETERMINE RELEVANT KEYWORDS
STEP S5
CALCULATE CORRELATIONS OF KEYWORDS AND DOCUMENTS AND DETERMINE CORRELATED DOCUMENTS
STEP S6
GENERATE CHARACTER STRING GRAPH AND SHRINK GRAPH
STEP S7
GENERATE INDEX HYPERTEXT
STEP S8
END

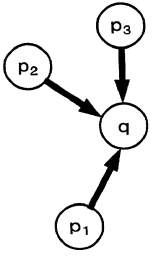
FIG. 7



CIRCLE(○) : WEB PAGE
THICKNESS OF ARROW(→) : LINK WEIGHT
PATTERN OF CIRCLE(○) : URL SIMILARITY

FIG. 8

$sim(p_i, q)=1$

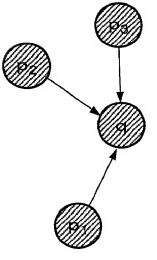


$$lw(p_i, q)=\frac{1}{sim(p_i, q)}=1$$

$$w_q=c_q+w_{p1}+w_{p2}+w_{p3}$$

FIG. 9A

$sim(p_i, q)=n+1$



$$lw(p_i, q)=\frac{1}{sim(p_i, q)}=\frac{1}{n+1}$$

$$w_q=c_q+\frac{w_{p1}+w_{p2}+w_{p3}}{n+1}$$

FIG. 9B

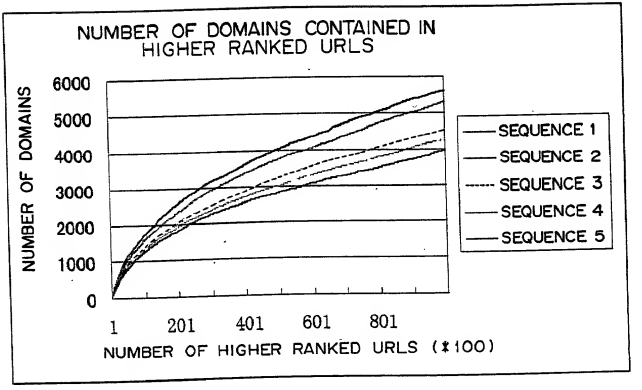

NUMBER OF DOMAINS CONTAINED IN HIGHER RANKED URLS
NUMBER OF DOMAINS
6000
5000
4000
3000
2000
1000
0
1
201
401
601
801
NUMBER OF HIGHER RANKED URLS (*100)
SEQUENCE 1
SEQUENCE 2
SEQUENCE 3
SEQUENCE 4
SEQUENCE 5

FIG. 10

Translation of EP 0 629 964 A1

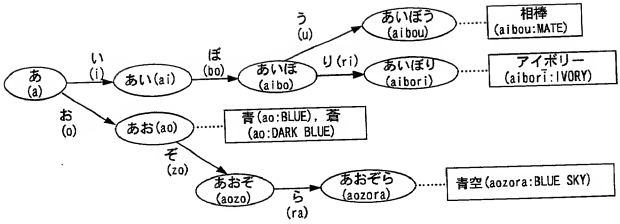


FIG. 11A

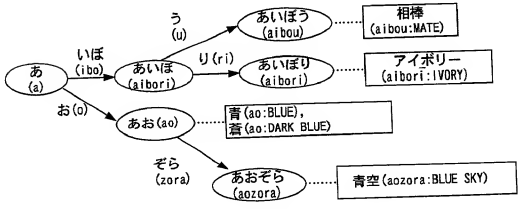


FIG. 11B

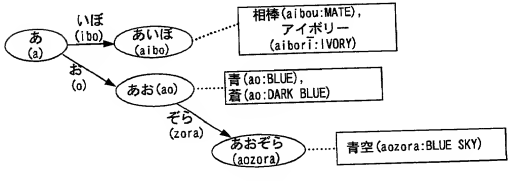


FIG. 11C

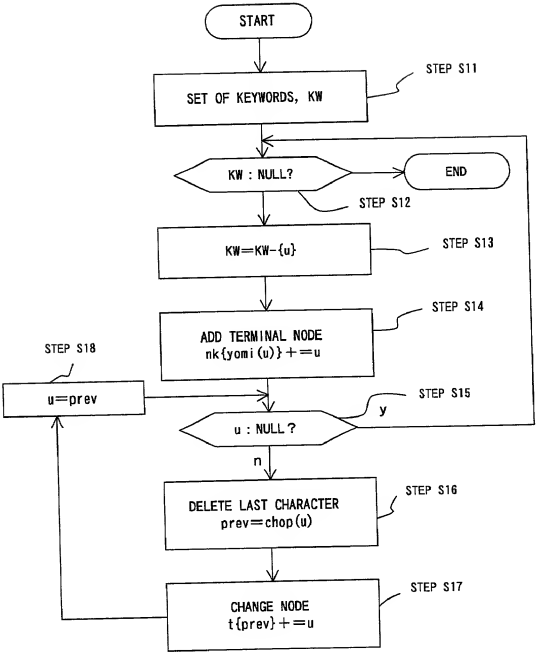

START
SET OF KEYWORDS, KW
STEP S11
KW : NULL?
END
STEP S12
KW=KW-{u}
STEP S13
ADD TERMINAL NODE
nk{yomi(u)}+=u
STEP S14
STEP S18
u=prev
STEP S15
y
u : NULL?
n
DELETE LAST CHARACTER
prev=chop(u)
STEP S16
CHANGE NODE
t{prev}+=u
STEP S17

FIG. 12

TOKUTO 2008-260

```
proc init_kw_graph ()
{
  @KW:set of keywords;      # SET OF KEYWORDS
  yomi : YOMI/Spell of keywords; # FUNCTION OR ARRAY THAT RETURNS PRONUNCIATION CHARACTERS OF KEYWORD
  foreach u in KW {          # FOR EACH KEYWORD u
   nk{yomi{u}} .= u."+";   # DESIGNATE nk() OF NODE OF PRONUNCIATION CHARACTERS OF KEYWORD u
   for ( i=0; i<length(u); i++) {    # REPEAT FOR LENGTH OF CHARACTER STRING OF KEYWORD u
        local prev = chop(u);          # DELETE LAST CHARACTER OF KEYWORD u AND ADD TO PARENT NODE
      t{prev} .= u."+";
        u = prev;
     }
  }
 }
```

FIG. 13

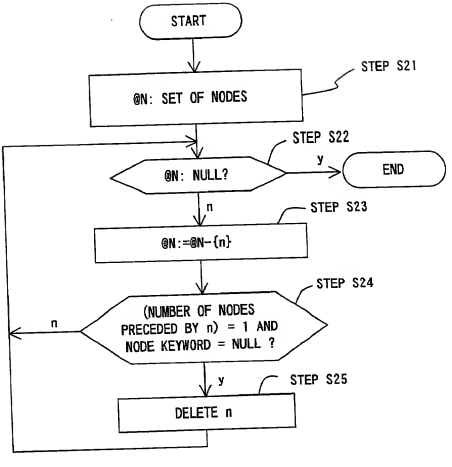

START
STEP S21
@N: SET OF NODES
STEP S22
@N: NULL?
y
END
n
STEP S23
@N:=@N-{n}
STEP S24
(NUMBER OF NODES
PRECEDED BY n) = 1 AND
NODE KEYWORD = NULL ?
n
y
STEP S25
DELETE n

FIG. 14

```
proc shrink_middle ()
{
  @N : set of nodes
  foreach n (@N) {
    next = t{n};     # NEXT NODE LIST
      kw = nk{n};      # KEYWORD LIST
      if (length(next) ==1 && kw == "") {
        delete(n)              # DELETE NODE n
      }
    }
}
```

FIG. 15

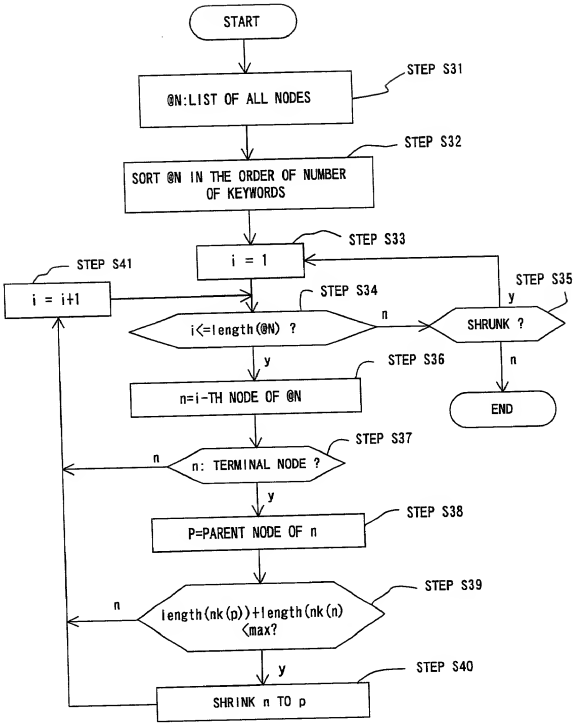

START
@N:LIST OF ALL NODES
STEP S31
SORT @N IN THE ORDER OF NUMBER OF KEYWORDS
STEP S32
i = 1
STEP S33
i = i+1
STEP S41
i<=length(@N) ?
STEP S34
n
y
SHRUNK ?
STEP S35
y
n
END
n=i-TH NODE OF @N
STEP S36
n: TERMINAL NODE ?
STEP S37
n
y
P=PARENT NODE OF n
STEP S38
length(nk(p))+length(nk(n) <max?
STEP S39
n
y
SHRINK n TO p
STEP S40

FIG. 16

```
proc shrink_leaf ()
{
  @N: set of nodes;      # NODE LIST
  word_max=2;              # word_max : IN THIS EXAMPLE, 2
  changed=true;            # WHEN KEYWORD IS TRANSFERRED, true
  @N= sort by_nk_length @N;  # SORTING IN ASCENDING ORDER OF NUMBER OF KEYWORDS
  while (changed) {        # CONTINUING WHILE TRANSFER IS PERFORMED
      changed=false;
      foreach n in @N {
       if (is_leaf(n)) {      # IN THE CASE OF TERMINAL NODE
          p=parent_node(n);       # PARENT NODE
          if (length(nk{p}) + length(nk{n}) <word_max) {
             nk{p}.=nk{n}."+";  # TRANSFERRING KEYWORD
             delete (n);      # DELETE TERMINAL NODE
             changed =true;   # PROOF OF TRANSFER
          }
         }
      }
    }
}
```

FIG. 17

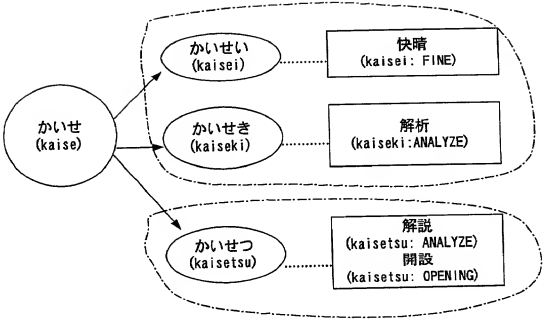

かいせ
(kaise)
かいせい
(kaisei)
快晴
(kaisei: FINE)
かいせき
(kaiseki)
解析
(kaiseki:ANALYZE)
かいせつ
(kaisetsu)
解説
(kaisetsu: ANALYZE)
開設
(kaisetsu: OPENING)

FIG. 18

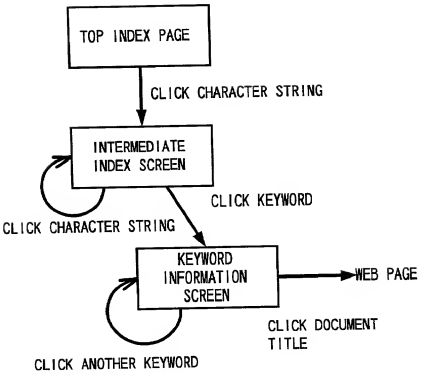

TOP INDEX PAGE
CLICK CHARACTER STRING
INTERMEDIATE INDEX SCREEN
CLICK KEYWORD
CLICK CHARACTER STRING
KEYWORD INFORMATION SCREEN
WEB PAGE
CLICK DOCUMENT TITLE
CLICK ANOTHER KEYWORD

FIG. 19

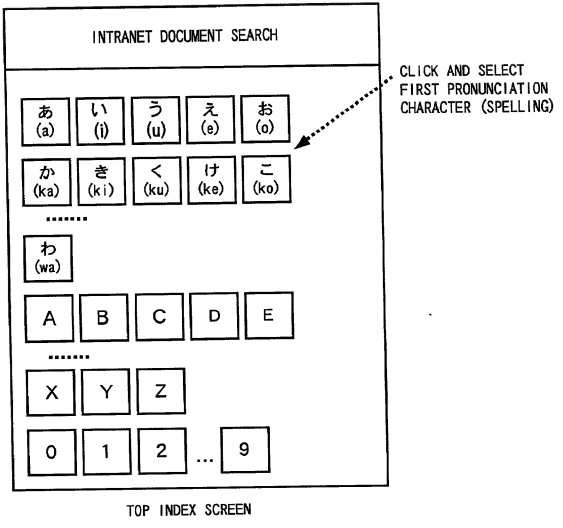

INTRANET DOCUMENT SEARCH
CLICK AND SELECT FIRST PRONUNCIATION CHARACTER (SPELLING)
あ (a)
い (i)
う (u)
え (e)
お (o)
か (ka)
き (ki)
く (ku)
け (ke)
こ (ko)
わ (wa)
A
B
C
D
E
X
Y
Z
0
1
2
9
TOP INDEX SCREEN

FIG. 20

TOKETO'29089760

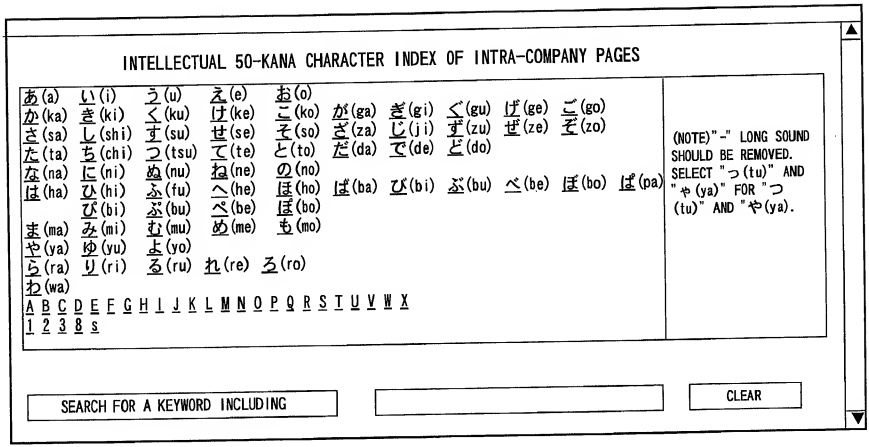


FIG. 21

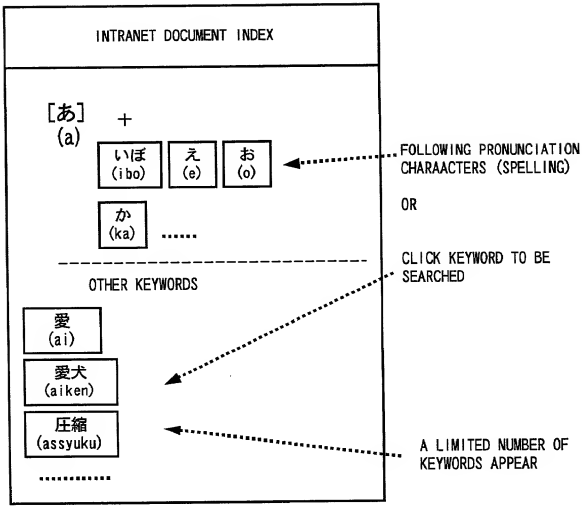

INTRANET DOCUMENT INDEX
[あ]
(a)
+
いぼ
(ibo)
え
(e)
お
(o)
か
(ka)
.....
FOLLOWING PRONUNCIATION CHARAACTERS (SPELLING)
OR
CLICK KEYWORD TO BE SEARCHED
OTHER KEYWORDS
愛
(ai)
愛犬
(aiken)
圧縮
(assyuku)
.....
A LIMITED NUMBER OF KEYWORDS APPEAR

FIG. 22

TOHZTO' 29089260

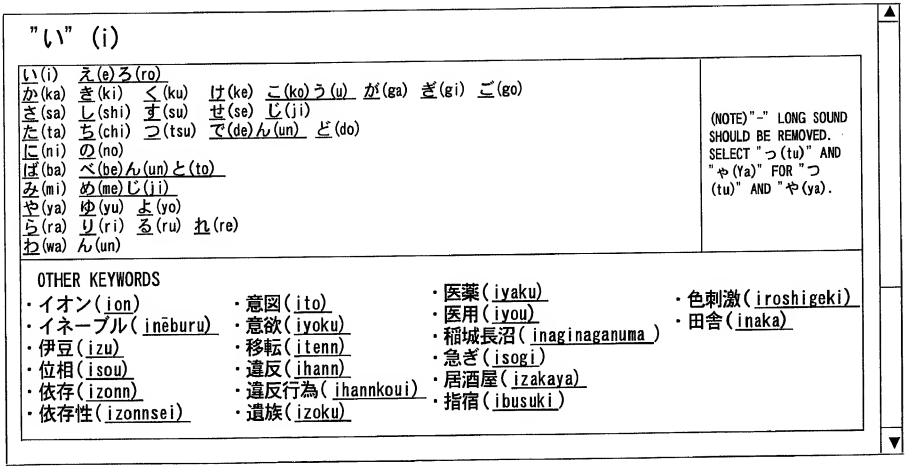


FIG. 23

TOMZTD-29089260

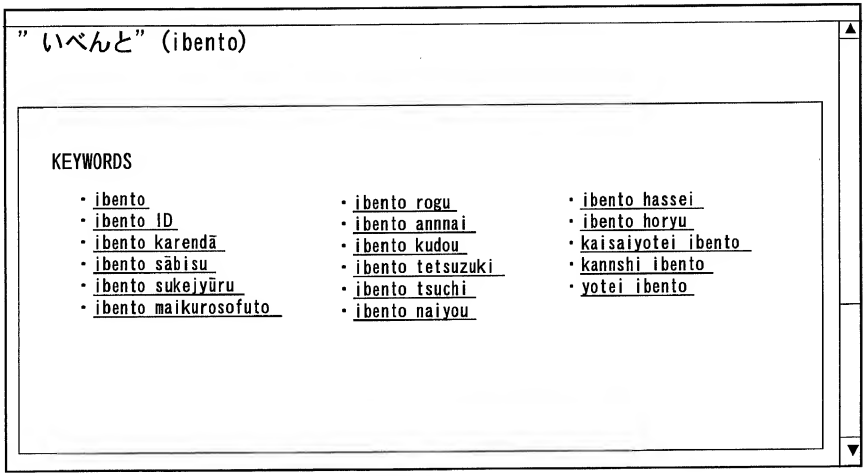


FIG. 24

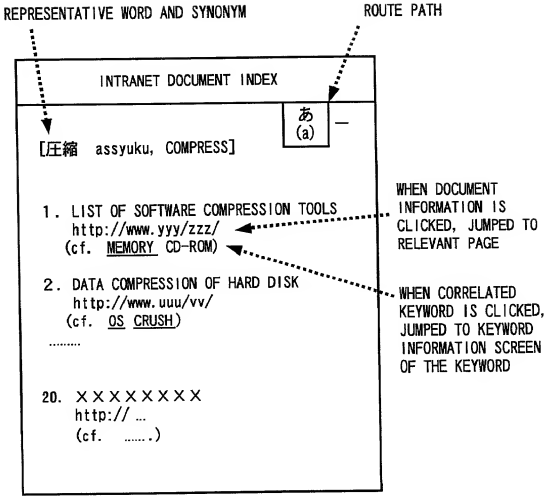

REPRESENTATIVE WORD AND SYNONYM
ROUTE PATH
INTRANET DOCUMENT INDEX
あ
(a)
[圧縮 assyuku, COMPRESS]
1. LIST OF SOFTWARE COMPRESSION TOOLS
http://www.yyy/zzz/
(cf. MEMORY CD-ROM)
2. DATA COMPRESSION OF HARD DISK
http://www.uuu/vv/
(cf. OS CRUSH)
.....
20. XXXXXXXX
http:// ...
(cf. .....)
WHEN DOCUMENT INFORMATION IS CLICKED, JUMPED TO RELEVANT PAGE
WHEN CORRELATED KEYWORD IS CLICKED, JUMPED TO KEYWORD INFORMATION SCREEN OF THE KEYWORD

FIG. 25

EP 0 762 290 A2

FIG. 26

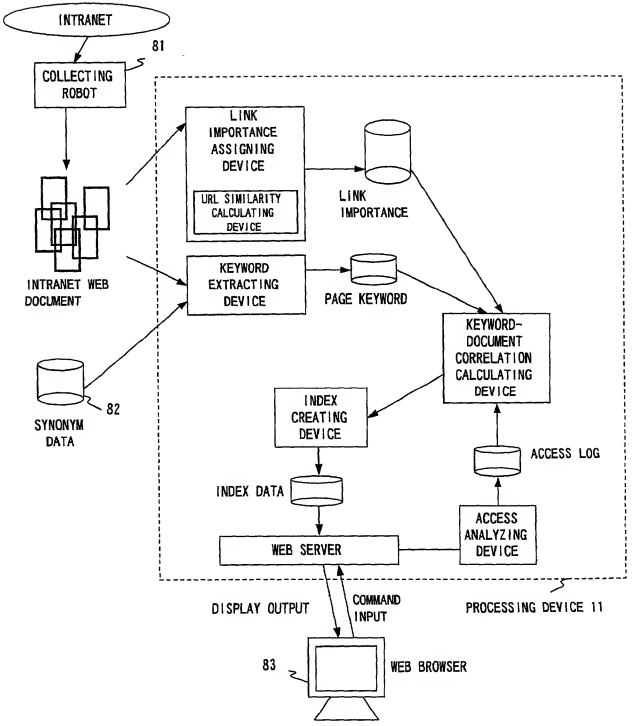

INTRANET
81
COLLECTING ROBOT
LINK IMPORTANCE ASSIGNING DEVICE
URL SIMILARITY CALCULATING DEVICE
LINK IMPORTANCE
INTRANET WEB DOCUMENT
KEYWORD EXTRACTING DEVICE
PAGE KEYWORD
KEYWORD-DOCUMENT CORRELATION CALCULATING DEVICE
SYNONYM DATA
82
INDEX CREATING DEVICE
ACCESS LOG
INDEX DATA
ACCESS ANALYZING DEVICE
WEB SERVER
DISPLAY OUTPUT
COMMAND INPUT
PROCESSING DEVICE 11
83
WEB BROWSER

FIG. 27

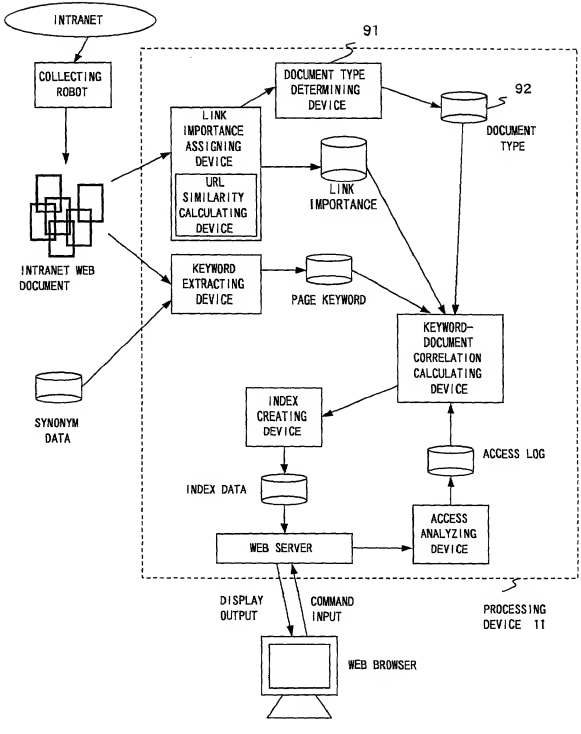

INTRANET
COLLECTING ROBOT
91
DOCUMENT TYPE DETERMINING DEVICE
92
DOCUMENT TYPE
LINK IMPORTANCE ASSIGNING DEVICE
URL SIMILARITY CALCULATING DEVICE
LINK IMPORTANCE
INTRANET WEB DOCUMENT
KEYWORD EXTRACTING DEVICE
PAGE KEYWORD
KEYWORD-DOCUMENT CORRELATION CALCULATING DEVICE
SYNONYM DATA
INDEX CREATING DEVICE
ACCESS LOG
INDEX DATA
ACCESS ANALYZING DEVICE
WEB SERVER
DISPLAY OUTPUT
COMMAND INPUT
PROCESSING DEVICE 11
WEB BROWSER

FIG. 28

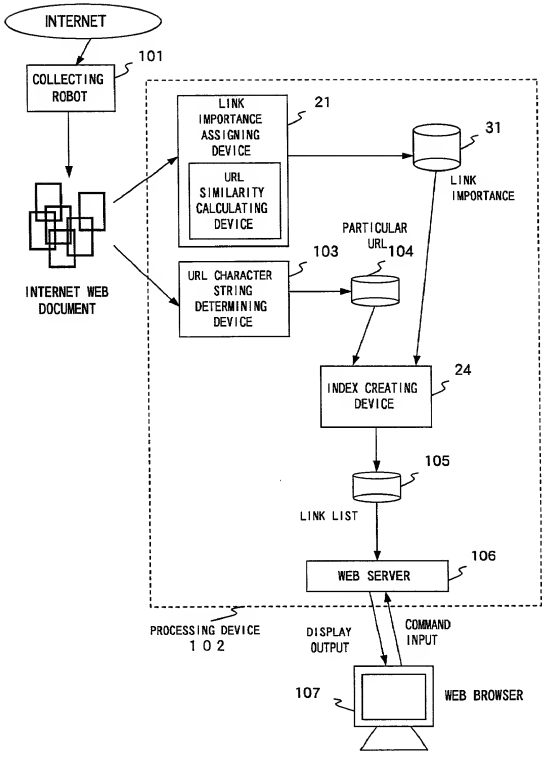

INTERNET
101
COLLECTING ROBOT
LINK IMPORTANCE ASSIGNING DEVICE
21
URL SIMILARITY CALCULATING DEVICE
31
LINK IMPORTANCE
INTERNET WEB DOCUMENT
PARTICULAR URL
103
104
URL CHARACTER STRING DETERMINING DEVICE
24
INDEX CREATING DEVICE
105
LINK LIST
106
WEB SERVER
PROCESSING DEVICE
102
DISPLAY OUTPUT
COMMAND INPUT
107
WEB BROWSER

FIG. 29

Printed 2/5/2004

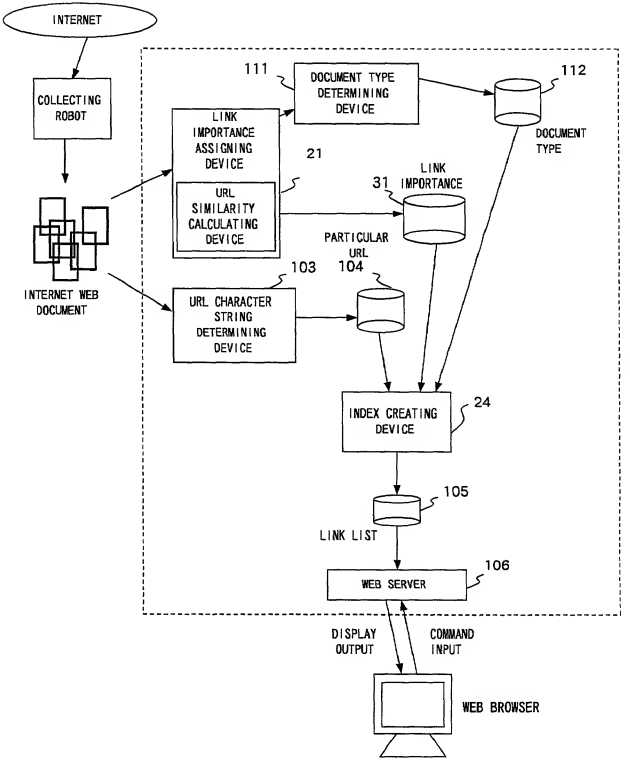


FIG. 30

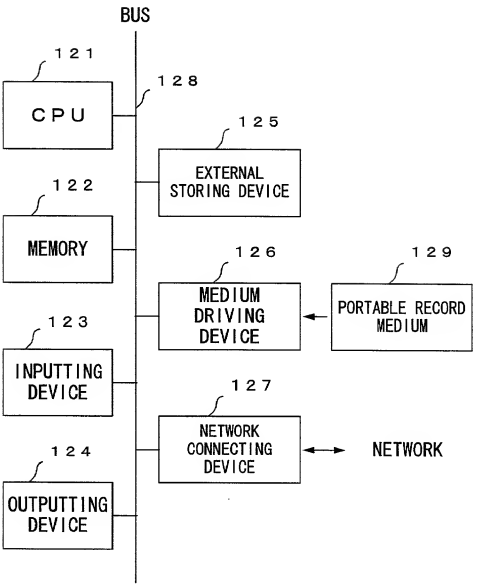

BUS
121
CPU
128
125
EXTERNAL STORING DEVICE
122
MEMORY
126
129
MEDIUM DRIVING DEVICE
PORTABLE RECORD MEDIUM
123
INPUTTING DEVICE
127
NETWORK CONNECTING DEVICE
NETWORK
124
OUTPUTTING DEVICE

FIG. 31

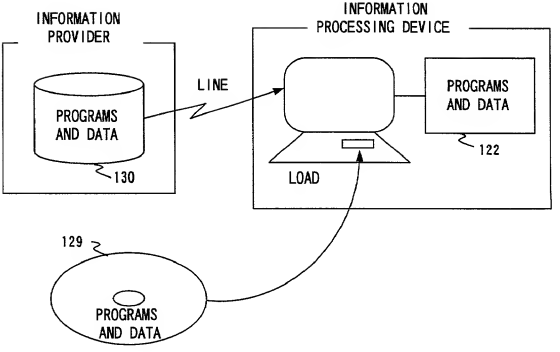

INFORMATION
PROVIDER
PROGRAMS
AND DATA
130
LINE
INFORMATION
PROCESSING DEVICE
PROGRAMS
AND DATA
122
LOAD
129
PROGRAMS
AND DATA

FIG. 32